Received: 22 January 2010 ■ *Revised: 4 June 2010* ■ *Accepted: 16 June 2010* ■ *Publication: 28 June 2010*

JNIRS JOURNAL OF NEAR INFRARED SPECTROSCOPY

Determination of olive oil adulteration with vegetable oils by near infrared spectroscopy coupled with multivariate calibration

Betül Öztürk, Ayşegül Yalçın and Durmuş Özdemir

Department of Chemistry, Faculty of Science, Izmir Institute of Technology, Gülbahçe 35340 URLA/İzmir, Turkey.
E-mail: durmusozdemir@iyte.edu.tr

There has been growing public awareness about the health benefits of olive oil throughout the world in recent years, resulting in a significant increase in its consumption as part of the daily diet. This demand has attracted fraudulent attempts to market olive oil which has been adulterated with cheaper oils. This study focuses on the near infrared (NIR) spectroscopic determination of adulteration of olive oil by vegetable oils using multivariate calibration. The binary, ternary and quaternary mixtures of olive, soybean, cotton, corn, canola and sunflower oils were prepared using a random design. The absorbance spectra of these synthetic samples were measured by a near infrared (NIR) spectrometer. A genetic algorithm-based variable selection algorithm, coupled with an inverse least squares multivariate calibration method (GILS) was used to build calibration models for possible adulterants and olive oil in the adulterated mixtures. The correlation coefficients of actual versus predicted concentrations resulting from multivariate calibration models for the different oils were between 0.90 and 0.99. The results demonstrated that NIR spectroscopy in conjunction with the GILS method makes it possible to determine the adulteration of olive oils regardless of adulterant vegetable oils over a wide range of concentrations.

Keywords: olive oil adulteration, near infrared spectroscopy, multivariate calibration, genetic algorithms, vegetable oils

Introduction

The commercial value of olive oil is considerably higher than other vegetable oils such as sunflower and corn oil. On account of increased public awareness about the health and nutritional benefits of olive oil in recent years, the adulteration of olive oil with relatively cheap edible oils becomes economically attractive. As a result, the determination and detection of adulterants in olive oils has become a very important issue for food safety and protection of consumers. A number of fraudulent attempts at marketing adulterated olive oils have stimulated regulatory authorities and researchers to develop fast and reliable analysis methods. Among these, chromatographic analysis permits the detection of adulterants in olive oils even at very low levels. In a detailed study about olive oil adulteration with sunflower, corn, peanut and coconut oils, gas chromatography (GC) coupled with mass spectrometry (MS) was used for the analysis of the lipid fractions.[1] Both qualitative and quantitative analyses were performed in order to determine both the type and the amount of the adulterant in olive oil samples and it was reported that the prediction ability of the models were around 90%. Recently, a review of the use of phytosterols as a tool for detection of olive oil adulteration by hazelnut oil was presented using GC and GC-MS[2] which focused on the lower detection limits of adulterants in olive oils, reported to be as low as 2% by mass. A comparative study about differentiating virgin olive oil from olive oil samples adulterated with vegetable oils based on volatile composition was reported.[3] Several chromatographic methods were widely used for the determination of olive oil adulteration.[4–6]

ISSN: 0967-0335
doi: 10.1255/jnirs.879

In addition, various physical and chemical analytical methods used for the purpose of establishing the authenticity of olive oil and detection of adulterants[7–9] have been reported.

Although chromatographic methods offer high sensitivity and accuracy, they are somewhat time-consuming and relatively expensive when compared with spectroscopic methods which offer faster and cheaper analysis.[10–14] Near infrared (NIR) spectroscopy finds widespread use in many different fields such as process monitoring,[15] biotechnology[16] and in the pharmaceutical and food industries.[17] Determination of olive oil adulteration with vegetable oils using NIR spectroscopy coupled with multivariate calibration has been reported in a number of previous studies.[18–23] Because NIR offers on-line, non-destructive and non-invasive determination of concentrations in multi-component mixture systems, it has become very popular for simultaneous chemical analysis. NIR spectroscopy is based on the absorption bands observed within a range between 780 nm and 2500 nm, which arise from overtones and combinations of fundamental mid-IR molecular vibrational bands. All of the vibrational modes can produce overtones but the most commonly observed bands arise from the C–H, O–H or N–H bonds in molecules.

Modern spectroscopic techniques offer fast analysis that can generate hundreds of spectra in a short period of time for multi-component samples. However, univariate calibration techniques fail to produce optimum results for those types of data as the components of the mixtures generally produce severely overlapping signals. Cases like this require a multivariate calibration approach in which instrumental responses measured at multiple wavelengths are related to a chemical or physical property of a sample even though it contains multiple components. In the last couple of decades, chemometrics and advanced computer technology have resulted in the development of several multivariate calibration techniques.[24–27]

Inverse least squares (ILS) is a multivariate calibration method based on the inverse Beer's law in which the concentration of an analyte is modelled as a function of absorbance. On the other hand, full spectral information collected from a spectroscopic technique produces hundreds of data points, if not thousands, for a given sample and often this spectrum contains many regions with collinearities and some amount of noise. In addition, it also contains absorbance regions in which the signal is not exactly linearly related to the concentration of the component being modelled. In these cases, ILS may not offer efficient solutions for a given problem with whole spectral information and, therefore, it might be necessary to apply a variable selection before the modelling step. Among several methods of variable selection, genetic algorithms (GAs) offer fast and efficient solutions for certain problems.[28–33] Genetic inverse least squares (GILS) is a modified version of the ILS method in which a small set of wavelengths is selected from a full spectral data matrix and evolved to an optimum solution using a GA. A detailed description of the GILS algorithm has been given in a number of reports elsewhere.[34–36]

In an earlier study, we applied NIR spectroscopy for the determination of olive oil adulteration in ternary mixtures of olive oil with corn and sunflower oil.[35] This study focuses on the determination of olive oil adulteration with other vegetable oils namely soybean, canola and cotton oil in addition to corn and sunflower oil in binary, ternary and quaternary mixtures of these oils using NIR spectroscopy in conjunction with genetic algorithm based multivariate calibration. The aim was to develop robust calibration models for the components of the mixtures regardless of the type of adulterant and the number of adulterants being used in a possible attempted fraud. The GILS method was used to establish calibration models with high predictive ability for the components of oil mixtures. In addition, selectivity of GA used in GILS was also investigated by obtaining the frequency distribution of selected wavelengths over the whole spectra of the samples.

Experimental section

Olive oil and other vegetable oils (corn, sunflower, canola, soybean and cotton) were obtained from local grocery stores. Mixtures of vegetable oils with olive oil were prepared in binary, ternary and quaternary mixtures using a random design giving a total of 160 samples. Concentration profiles of the samples in binary, ternary and quaternary mixtures of olive, corn, sunflower, canola, cotton and soybean oils are given in Tables 1–3. Among these multicomponent mixtures, binary sets of olive and cotton oils and olive and soybean oils contained 30 samples each as given in Table 1, whereas ternary mixtures of olive, cotton and soybean oils were formed with 50 samples as shown in Table 2. The quaternary mixture set formed from olive, sunflower, corn and canola oils which contained 50 samples is given in Table 3. Synthetic mixtures were prepared with a random design strategy in which concentrations of olive, corn, sunflower and soybean oils ranged from 0.0% to 100.0%, whereas the concentration of soybean and cotton oil were kept at 40.0% by mass in the mixtures. Because of possible correlation problems, an additional independent validation set with 30 samples was also prepared from olive, corn, sunflower and canola oils in such a way that one component concentration was kept constant while others were ranged between 10% and 50% by mass. Concentration profiles of these samples are given in the Results and Discussion section along with the predicted concentrations by GILS.

Spectra were collected using a Bio-Rad Excalibur FTS 3000 NX Fourier Transform Near Infrared spectrometer (Bio-Rad Laboratories Europe Ltd, Hemel Hempstead, Herts, UK) between 4000 cm^{-1} and 10,000 cm^{-1} with a resolution of 16 cm^{-1}. This spectrometer was equipped with a 250 W tungsten–halogen source, a calcium fluoride beamsplitter and a lead selenide detector. The measurements were carried out using 5 mm pathlength infrasil quartz sample holders (Starna, Atascadero, CA, USA). All spectra were then transferred to a computer for data processing. The GILS method was written using MATLAB programming language (Matlab 5.3; MathWorks Inc., Natick, MA, USA).

Table 1. Concentration profiles of binary mixtures of olive and cotton oils and olive and soybean oils given on a mass percent basis (w/w%).

No.	Olive	Cotton	No.	Olive	Soybean
1	86.87	13.13	1	87.01	12.99
2	66.04	33.96	2	65.97	34.03
3	88.03	11.97	3	99.00	1.00
4	98.82	1.18	4	95.88	4.12
5	95.96	4.04	5	68.98	31.02
6	64.08	35.92	6	81.00	19.00
7	68.98	31.02	7	81.99	18.01
8	80.99	19.01	8	61.96	38.04
9	93.01	6.99	9	59.93	40.07
10	81.90	18.10	10	75.84	24.16
11	62.05	37.95	11	84.96	15.04
12	74.00	26.00	12	69.95	30.05
13	60.02	39.98	13	83.18	16.82
14	76.04	23.96	14	65.99	34.01
15	90.92	9.08	15	89.95	10.05
16	84.91	15.09	16	75.97	24.03
17	69.98	30.02	17	66.99	33.01
18	73.96	26.04	18	97.99	2.01
19	83.32	16.68	19	62.99	37.01
20	66.00	34.00	20	95.91	4.09
21	80.00	20.00	21	60.97	39.03
22	89.98	10.02	22	87.82	12.18
23	75.94	24.06	23	64.01	35.99
24	67.92	32.08	24	92.99	7.01
25	67.01	32.99	25	73.95	26.05
26	97.97	2.03	26	90.86	9.14
27	96.97	3.03	27	73.97	26.03
28	63.00	37.00	28	80.00	20.00
29	95.85	4.15	29	67.95	32.05
30	81.97	18.03	30	96.91	3.09

Results and discussion

Most of the olive oil adulteration studies using NIR spectroscopy[18–20] reported in the literature have been concerned with two- or three-component mixtures. This results in some limitations in the use of the calibration models as the type and the number of adulterants in a real sample would be different from one case to another. In fact, there would be no prior information about the nature of the adulterant in a given unknown sample and, thus, a successful calibration model should be able to predict at least the amount of olive oil in unknown samples regardless of the other components in the suspected sample. Therefore, this study aimed at determining the amount of olive oil in several binary, ternary and quaternary mixtures of olive oil with various vegetable oils, namely corn, sunflower, canola, soybean and cotton oils. Mixtures that contained more than four components were also prepared; however, it became apparent that the success of the models was greatly reduced when samples containing more than four components were introduced into the calibration step. The reason for this can be seen when pure component spectra of these oils were examined. Figure 1 shows the NIR spectra of pure olive oil along with the vegetable oils used in this study.

The pure component spectra of all the vegetable oils were quite similar to the olive oil spectrum. Only minor differences in spectral absorbance values may be seen on closer examination of the frequency range between 4800 cm^{-1} and 4500 cm^{-1}. Throughout the multivariate calibration, these slight changes in spectral characteristics were used to generate individual calibration models for the components of the multi-component mixtures. Multivariate calibration models for each of the

Table 2. Concentration profiles of ternary mixtures of olive, cotton and soybean oils given on a mass percent basis (w/w%).

No.	Olive	Cotton	Soybean	No.	Olive	Cotton	Soybean
1	59.87	29.06	11.07	26	75.96	0.00	24.04
2	65.97	26.98	7.05	27	68.93	15.01	16.06
3	77.99	3.97	18.04	28	80.92	13.00	6.08
4	61.85	22.08	16.07	29	73.89	1.04	25.07
5	74.86	16.05	9.09	30	89.86	4.05	6.09
6	90.98	2.05	6.96	31	69.93	17.02	13.05
7	90.90	3.03	6.07	32	73.91	18.00	8.09
8	88.00	4.00	8.00	33	70.96	20.07	8.97
9	59.93	15.07	25.00	34	78.94	18.04	3.01
10	80.93	8.01	11.05	35	73.09	11.96	14.95
11	68.06	22.99	8.95	36	97.99	1.01	1.00
12	64.05	18.98	16.97	37	76.95	3.98	19.07
13	71.93	27.02	1.04	38	68.86	18.08	13.06
14	76.91	20.98	2.11	39	60.01	14.02	25.97
15	84.88	5.05	10.07	40	87.94	7.12	4.94
16	70.06	29.94	0.00	41	72.04	19.02	8.94
17	70.00	0.00	30.00	42	69.93	5.04	25.03
18	82.96	10.94	6.10	43	66.01	8.06	25.94
19	89.90	9.04	1.06	44	88.98	2.04	8.98
20	94.86	2.11	3.03	45	84.91	3.00	12.09
21	74.00	10.97	15.04	46	78.96	18.97	2.08
22	66.98	19.94	13.08	47	79.09	20.91	0.00
23	87.04	12.96	0.00	48	63.02	5.08	31.90
24	65.94	12.02	22.04	49	85.93	1.02	13.04
25	59.87	29.06	11.07	50	62.96	28.03	9.01

oils mentioned above were generated using a GILS method and the performance of these models was tested with independent prediction sets. For this, about one-third of the samples for each component were reserved for prediction and the remaining two-thirds of the samples were used as calibration sets. Because the GILS method is an iterative procedure, it was necessary to optimise calibration models with a cross-validation approach in order to avoid possible overfitting problems during the model-building step. Even though the calibration models were generated with cross-validation during iterations of the genetic algorithm, an independent prediction set would better determine the stability of the models in the prediction of the true unknown samples. The success of individual multivariate calibration models was evaluated using both the standard error of cross-validation (*SECV*) and the standard error of prediction (*SEP*) for the independent prediction set. As a result of the random nature of the GILS algorithm, it generated slightly different *SECV* results from one run to another, since the selected data points on the whole spectrum were different in each run. For this reason, the GILS algorithm was set to run 100 times with 50 iterations and 30 genes for each component of the mixtures. Among these several runs, the one with the lowest *SECV* was selected for further prediction of the independent samples. Figure 2 shows the actual versus predicted plots of all the oils used in this study.

As seen from Figure 2, the correlation coefficients of determination for the oils used in this study ranged from 0.90 for canola oil to 0.99 for olive oil. The plots given in Figure 2 were generated from the models that produced the lowest *SECV*. In general, it is expected that multivariate calibration with a cross- validation step will generate *SEP* values for the independent prediction set which are close to the *SECV*, but the actual versus predicted plots given in Figure 2 show that somewhat larger *SEP* values were apparent. Among the multivariate calibration models generated, olive oil and soybean oil models were the most successful for the prediction of the independent samples, as shown in Table 4, which summarises the *SECV* and *SEP* results for all of the oils.

Among the six different oils that were studied here, the multivariate calibration model for soybean has the lowest *SECV* and *SEP* values, whereas the largest values were obtained for canola oil. On the other hand, the *SECV* value for cotton oil seems quite low; however, the model resulted in a

Table 3. Concentration profiles of quaternary mixtures of olive, sunflower, corn and canola oils given on a mass percent basis (w/w%).

No.	Olive	Sunflower	Corn	Canola	No.	Olive	Sunflower	Corn	Canola
1	100.00	0.00	0.00	0.00	26	24.91	28.81	22.61	23.68
2	0.00	100.00	0.00	0.00	27	18.49	16.58	34.27	30.65
3	0.00	0.00	100.00	0.00	28	38.46	37.86	5.14	18.55
4	0.00	0.00	0.00	100.00	29	8.06	31.94	15.88	44.12
5	49.67	50.33	0.00	0.00	30	14.29	27.19	27.73	30.79
6	49.76	0.00	50.24	0.00	31	6.84	15.31	71.65	6.21
7	49.99	0.00	0.00	50.01	32	14.93	11.18	42.86	31.03
8	0.00	50.07	49.93	0.00	33	25.88	22.93	22.08	29.11
9	0.00	50.03	0.00	49.97	34	45.52	25.42	21.41	7.64
10	0.00	0.00	50.01	49.99	35	46.10	17.10	34.02	2.77
11	33.30	33.46	33.23	0.00	36	18.89	23.67	23.36	34.09
12	33.32	33.38	0.00	33.30	37	21.19	28.31	5.68	44.82
13	33.67	0.00	33.22	33.11	38	14.38	18.45	40.44	26.73
14	0.00	33.33	33.28	33.39	39	29.25	26.74	41.92	2.08
15	24.97	24.97	25.08	24.98	40	14.56	34.52	20.69	30.23
16	10.06	19.93	30.01	40.01	41	7.86	15.66	3.50	72.98
17	39.90	10.02	19.94	30.14	42	10.15	19.19	35.26	35.40
18	30.02	39.96	9.99	20.03	43	4.84	19.72	37.50	37.95
19	19.96	29.94	39.96	10.14	44	25.20	29.96	10.75	34.10
20	3.36	69.93	4.28	22.43	45	45.62	10.19	38.65	5.54
21	22.57	24.46	8.07	44.90	46	0.86	32.10	17.62	49.42
22	36.39	23.28	2.55	37.78	47	33.70	20.56	27.12	18.62
23	76.03	1.35	15.04	7.59	48	3.45	24.15	36.43	35.97
24	49.78	33.83	2.24	14.15	49	44.04	25.12	20.38	10.46
25	74.08	3.77	5.50	16.65	50	18.48	23.83	4.08	53.61

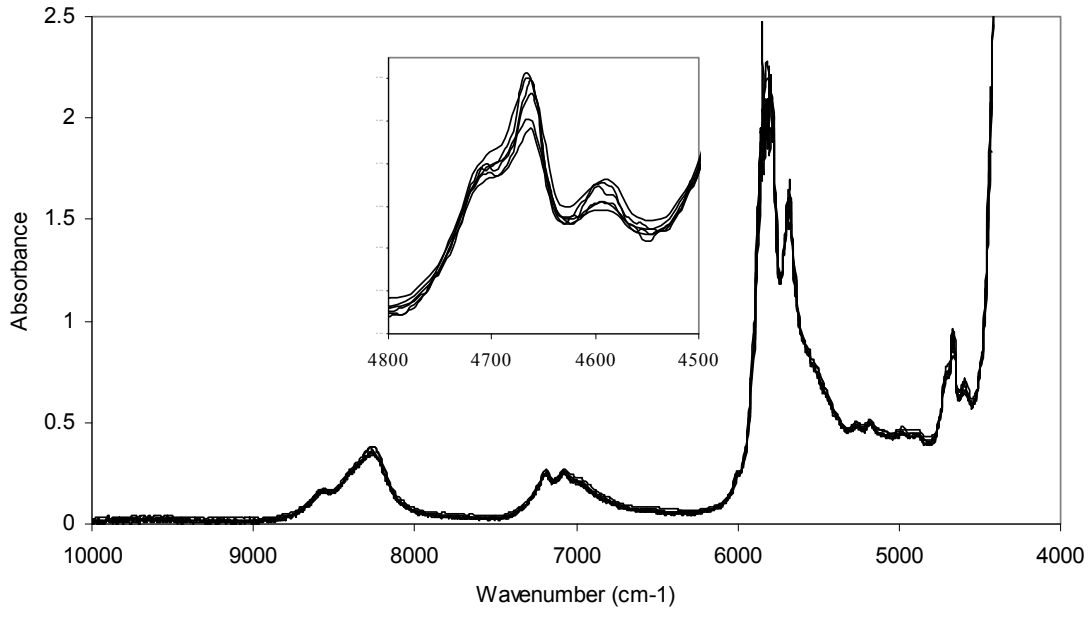

Figure 1. Pure component near infrared spectra of olive, corn, sunflower, canola, soybean and cotton oils.

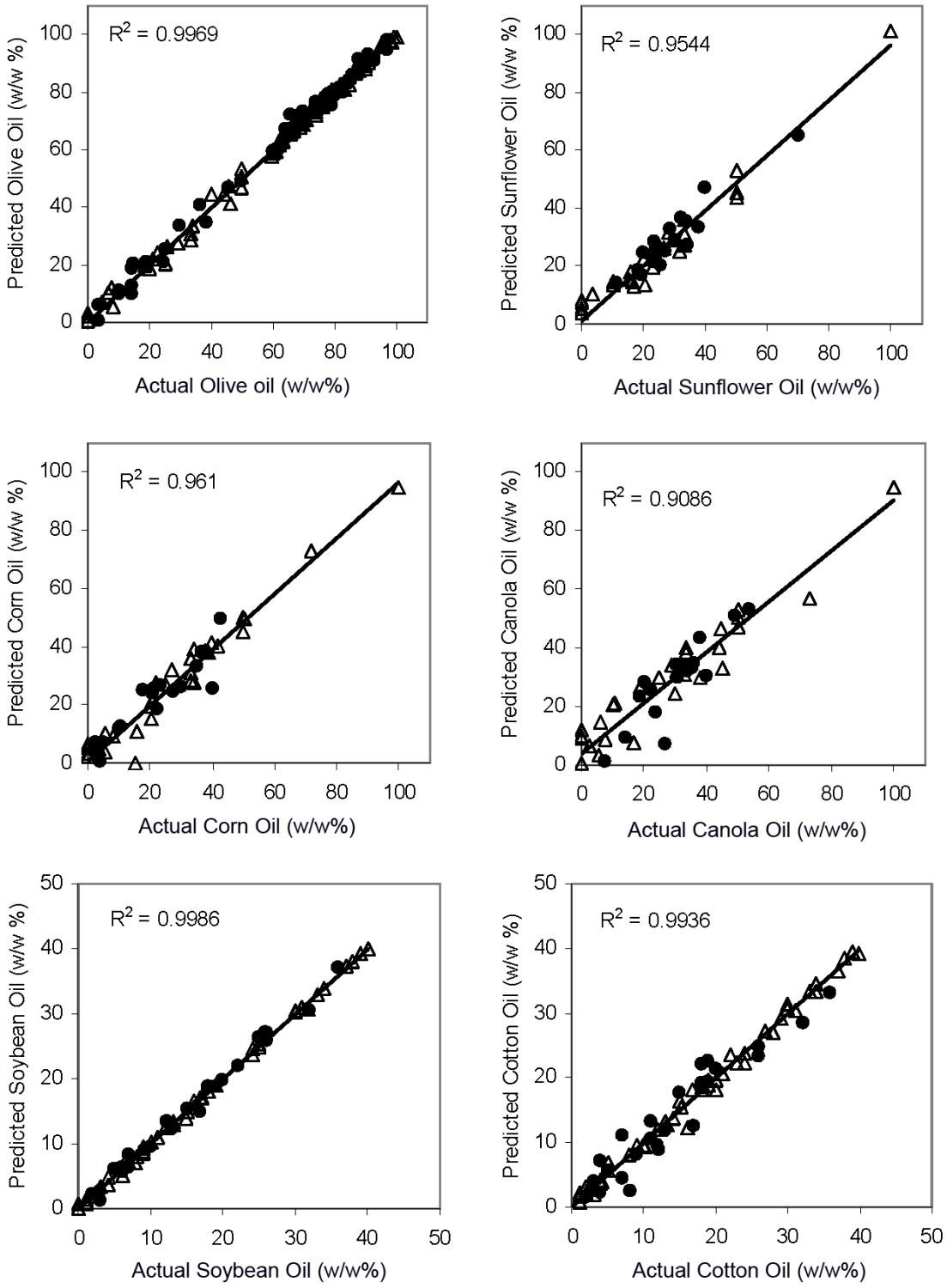

Figure 2. Actual content of olive oil and adulterant vegetable oils; corn, sunflower, canola, soybean and cotton oils versus GILS predicted plots. (△, calibration set samples, ●, validation set samples).

somewhat larger *SEP* value. In addition, relatively larger *SECV* and *SEP* values were observed for corn, sunflower and canola oils. In fact, the *SEP* values for sunflower and canola oil were smaller than the corresponding *SECV* values. The possible explanation for these results would be the number of samples that contained corn, sunflower and canola oils. There were only 50 samples that contained these components and multivariate calibration models were generated with two-thirds of

Table 4. Standard error of cross-validation (*SECV*) and standard error of prediction (*SEP*) results of components of the oil mixtures given on a mass percent basis (w/w%).

(w/w %)	Olive oil	Corn oil	Sunflower oil	Canola oil	Soybean oil	Cotton oil
SECV	1.58	4.71	4.58	7.60	0.45	0.96
SEP	2.45	5.09	3.93	6.41	0.94	2.81

these samples, whereas more than 50 samples were used to generate calibration samples for soybean and cotton oils and 106 samples were used for the olive oil model. It is also important to mention that soybean and cotton oils were only presented in binary and ternary mixtures with olive oil whereas corn, sunflower and canola oils were prepared in binary, ternary and quaternary mixtures with olive oil. Therefore, there was an increased complexity in their samples which resulted in less successful calibration models. However, it should be realised that most adulterations of olive oil involve the addition of the cheapest vegetable oil and generally one or two components. The primary concern in the determination of olive oil with other vegetable oils would be focus on quantitative determination of olive oil regardless of the type and amount of other vegetable oils. Therefore, the calibration model obtained for olive oil in this study requires more attention. As given in Table 4, GILS was able to predict the amount of olive oil with low prediction errors in various different compositions including binary, ternary and quaternary mixtures with several different vegetable oils.

Even though the GILS is involved in several random processes because of its iterative nature, the method is expected to concentrate on wavelength or wavenumber regions of the whole spectral range where the absorbances have highest correlation with analyte concentrations. To illustrate this, frequency distributions of the selected wavenumbers in 100 runs were plotted along with an NIR spectrum of pure component spectra in Figure 3 for each of the oils studied.

As may be seen from the frequency distribution of the selected wavenumbers, GILS focused on different wavenumber regions in order to build successful calibration models for each vegetable oil including olive oil. Overall, the most frequently selected and retained wavenumbers throughout iterative cycles of GILS do not seem to concentrate on specific narrow portions of whole spectra. However, there are regions in different parts of the spectrum where selection frequency reached a value of 40 and more. For example, the wavenumber region around 5900 cm^{-1} was the most frequently selected region for olive oil determination whereas the 4500 cm^{-1} region was preferred more often than any other part of the spectrum for soybean oil. When selection frequency distribution of the most retained wavenumber regions for corn oil was examined, it was seen that the GILS method was not able to concentrate on a narrow region but rather a wider wavenumber region between 10,000 cm^{-1} and 6500 cm^{-1}. As in the case of soybean oil, the wavenumber selection frequency distribution of cotton oil was also located in the 5100 cm^{-1} to 4500 cm^{-1} spectral region.

Because samples were prepared on a mass percent basis for the components of the oil mixtures, concentration of one component was related to the other components and this may have caused some correlation problems. In addition, it is of interest to see how the proposed GILS calibration method performed in comparison with a standard multivariate calibration methods such as partial least squares (PLS) regression. In order to take account of these issues, the additional olive, corn, sunflower and canola oil ($n = 30$) samples were analysed as described above and spectra were processed by both GILS and PLS methods. The concentrations of the quaternary mixtures are given in Table 5, along with predicted concentrations by both GILS and PLS methods; the number of PLS factors selected from leave-one-out cross-validation and corresponding *SEP* values are also included. Selection of the optimum number of PLS factors was done with prediction error sum of squares (*PRESS*) from leave-one-out cross-validation. It should be noted that GILS is not a factor-based calibration method; *SEP* of olive oil was 2.93% by mass using the GILS method. On the other hand, PLS predictions were relatively poor with the *SEP* values ranging from 4.64% to 8.30% by mass. As may be seen in Table 5, calibration models built with PLS required a relatively large number of PLS factors ranging from 11 to 15.

When the *SEP* values of GILS given in Table 5 were compared with values given in the Table 4, it is seen that comparable results were obtained. For example, concentration of olive oil was set to 10% in the first five samples while the corn and sunflower oil samples were varied between 10% to 50% and canola oil concentration kept at 30% by mass in order to have 100% total. As seen in the table, predicted concentration of olive oil in these samples ranged from as low as 3.15% to 14.16% by mass. The residuals of predicted olive, corn, sunflower and canola oils were plotted as a function of sample numbers in Figure 4.

In this figure, residuals for olive oil ranged between –5% and +5% by mass, whereas for the other vegetable oils larger prediction errors were reported. The magnitude of the residuals were about twice those for olive oil, ranging from –10% to +10%. Overall, these results for the second independent validation set indicated that GILS was relatively more successful in the prediction of olive oil in the quaternary mixtures of other vegetable oils. However, care must be taken that the residuals obtained from multivariate calibration models generated for corn, sunflower and canola oil were relatively large and the prediction errors tend to go up as the number of components increases in the multicomponent mixtures.

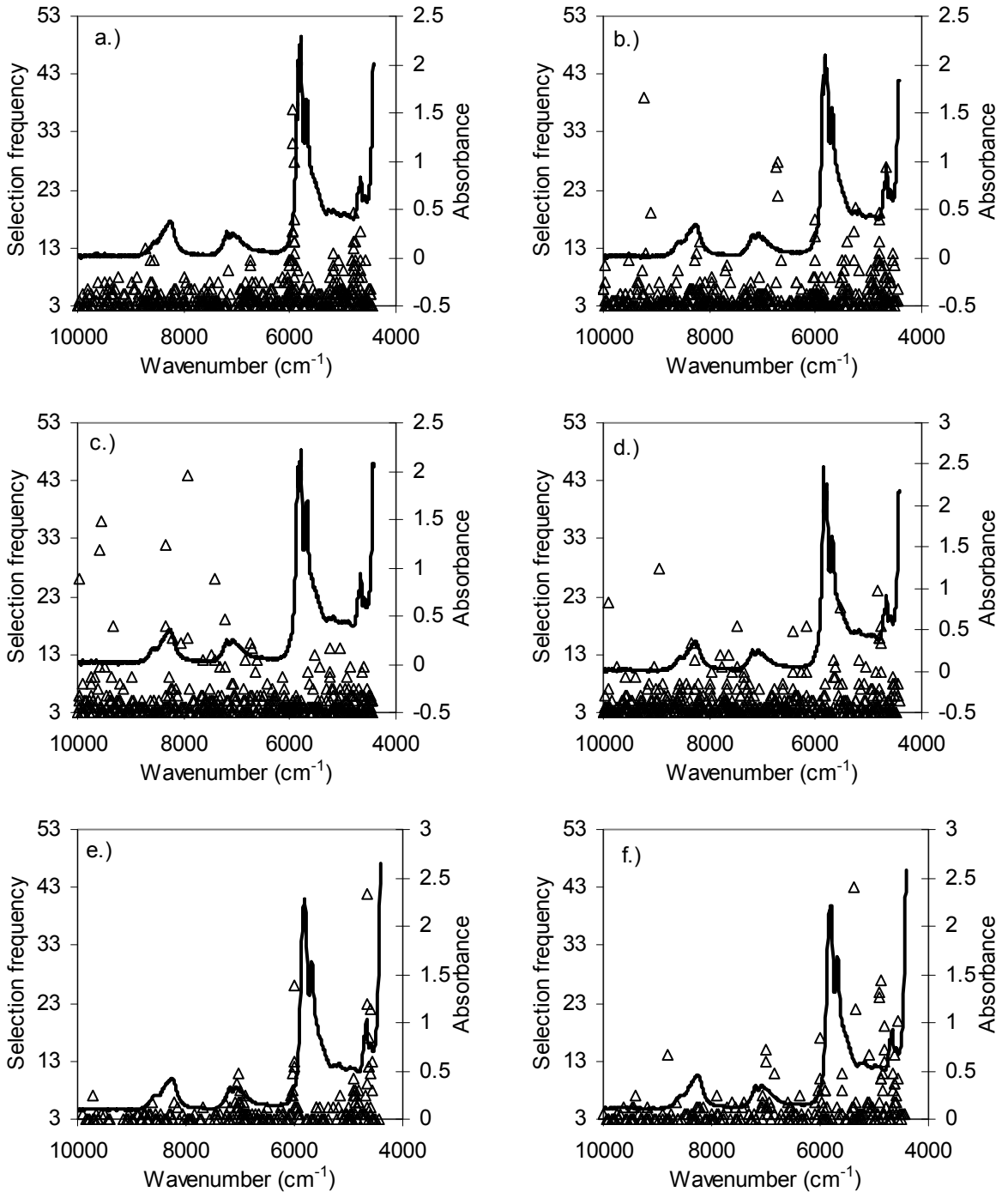

Figure 3. Frequency distributions of the selected wavenumbers by GILS method along with the corresponding pure component spectrum; (a) olive oil, (b) sunflower oil, (c) corn oil, (d) canola oil, (e) soybean oil and (f) cotton oil. (△, selection frequency, ——, pure component spectrum).

Conclusions

This study has demonstrated that NIR spectroscopy with multivariate calibration can be used to determine olive oil adulterations with various vegetable oils, regardless of the type and amount of the adulterant oil used in a possible fraudulent attempt. However, it is important to realise that standard error of prediction (*SEP*) values are all above 2% (w/w) except for soybean oil and, therefore, quantitative determination of minor amounts of (for example, < 5% by mass) adulterants would be question-

Table 5. Actual and predicted percent concentrations of olive, corn, sunflower and canola oil mixtures in the second independent validation set along with the number of PLS factors and corresponding standard error of prediction (*SEP*, w/w%) values from GILS and PLS.

No.	Olive oil (w/w%)			Corn oil (w/w%)			Sunflower oil (w/w%)			Canola oil (w/w%)		
	Actual	Predicted		Actual	Predicted		Actual	Predicted		Actual	Predicted	
		GILS	PLS		GILS	PLS		GILS	PLS		GILS	PLS
1	10.83	14.16	14.75	10.04	20.50	18.97	49.25	46.68	39.35	29.88	24.21	24.59
2	9.99	3.15	9.70	20.06	18.98	18.90	39.95	47.48	38.01	30.01	24.77	33.35
3	10.02	5.98	3.98	29.95	29.12	39.70	30.03	30.31	23.41	30.00	21.01	32.39
4	10.18	11.15	2.80	39.86	41.79	28.42	19.98	17.50	23.79	29.98	26.30	44.68
5	10.07	6.22	7.53	49.84	56.48	40.64	10.16	5.38	11.62	29.93	30.41	40.33
6	19.91	15.65	19.91	9.99	10.21	21.42	49.73	51.60	50.02	20.37	24.09	16.01
7	19.99	22.50	25.45	20.09	18.66	27.77	39.96	29.97	35.40	19.97	23.92	11.17
8	20.10	16.44	18.46	29.92	33.12	28.20	29.96	29.88	26.27	20.02	20.51	20.02
9	19.97	17.39	20.20	39.72	38.67	40.02	20.06	24.31	23.71	20.25	27.65	18.11
10	19.95	24.31	16.55	49.95	52.31	47.84	9.99	17.75	8.50	20.11	18.41	26.08
11	29.68	29.39	24.22	9.88	7.59	12.68	49.24	40.78	44.75	11.20	13.36	24.87
12	29.97	31.89	23.11	19.98	20.54	14.56	39.95	38.11	34.38	10.10	9.93	24.14
13	30.03	29.43	26.31	29.98	34.61	20.18	30.02	30.19	37.24	9.97	14.38	17.49
14	30.09	32.20	32.22	39.91	40.75	29.10	19.99	19.00	25.54	10.01	–1.32	13.93
15	29.96	33.37	33.83	49.93	56.43	38.25	10.02	7.97	15.87	10.09	19.58	6.78
16	10.02	9.17	19.78	10.20	11.95	19.29	29.97	14.46	26.90	49.81	49.80	31.33
17	20.08	19.49	22.27	10.13	12.19	17.28	29.98	31.86	30.52	39.81	35.15	32.76
18	29.87	29.21	28.27	10.05	10.86	3.22	30.08	30.30	34.19	30.00	35.29	31.62
19	39.85	40.69	37.89	10.19	2.90	8.90	29.96	28.83	32.47	20.00	16.84	23.76
20	49.66	50.01	46.24	10.55	21.15	26.39	29.93	29.88	11.82	9.86	16.91	16.48
21	10.11	9.14	16.70	19.81	20.90	18.05	20.48	19.83	25.03	49.60	57.48	40.88
22	20.33	18.99	24.74	19.90	20.12	24.74	19.92	20.04	16.84	39.86	52.48	32.60
23	29.99	29.98	24.59	19.98	10.04	21.06	20.06	14.16	20.36	29.98	34.06	34.43
24	39.67	37.14	37.14	20.56	12.98	18.62	19.94	27.32	16.53	19.83	14.11	23.49
25	50.00	45.44	48.52	19.93	14.74	17.36	20.04	20.35	15.62	10.03	9.71	13.25
26	10.12	6.61	19.49	29.81	29.12	24.27	10.01	12.74	23.61	50.06	58.96	34.71
27	19.76	23.73	26.32	29.62	32.03	25.14	10.36	16.24	20.14	40.27	38.08	26.96
28	29.49	32.26	32.58	29.52	32.78	36.43	10.70	5.33	12.70	30.29	28.35	23.54
29	39.93	42.16	38.18	29.88	24.63	33.82	10.31	10.15	7.03	19.88	28.34	23.04
30	49.05	52.28	44.99	30.86	26.82	19.74	10.10	3.26	20.51	9.99	12.81	12.33
PLS factors		—	14	—	—	11	—	—	13	—	—	15
SEP (w/w%)		2.93	4.64	—	4.71	7.49	—	5.19	6.34	—	5.86	8.30

able. Also, the predictive ability of the models decreases with increasing numbers of vegetable oils used for adulteration of olive oil. Nevertheless, it was concluded that multi-component mixtures containing up to four oils could be successfully modelled with a wide concentration range. In addition, the frequency distribution of most frequently selected wavenumbers showed that the GILS method was able to generate selective multivariate calibration models for the components of the mixtures using variable selection. The genetic algorithm used in GILS helped to optimise calibration models in order to extract relevant information for each component. When compared with GILS, it is seen that PLS generated relatively larger prediction errors for the second independent validation set.

Acknowledgement

The financial support in this work was supplied by Scientific and Technical Council of Turkey (TUBİTAK) through the project No: 107T037.

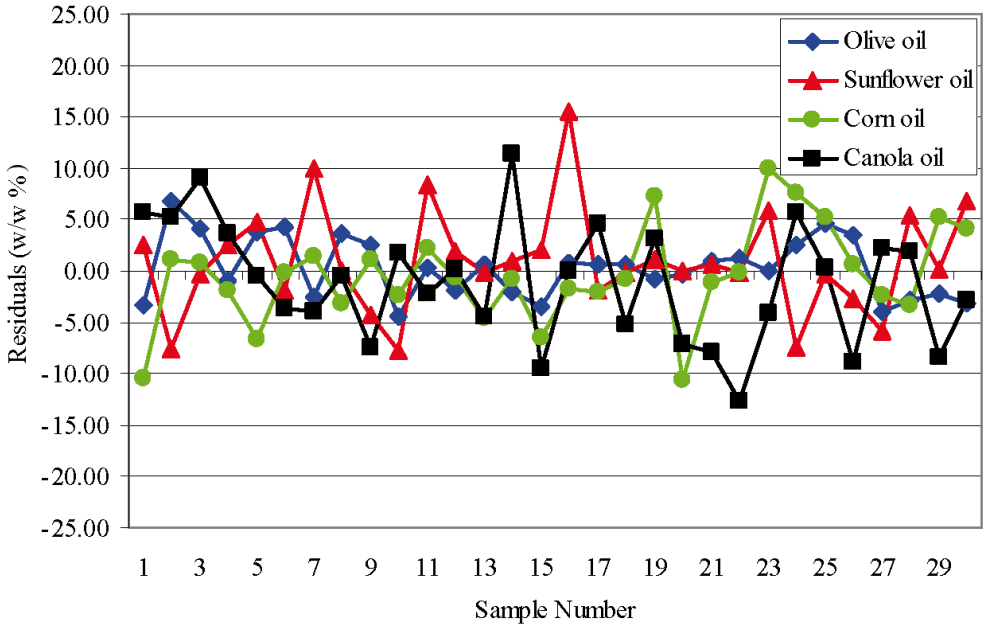

Figure 4. Residuals of predicted olive, corn, sunflower and canola oils as a function of sample number in the second validation set using GILS.

References

1. F. Priego Capote, J. Ruiz Jiménez and M.D. Luque de Castro, "Sequential (step-by-step) detection, identification and quantitation of extra virgin olive oil adulteration by chemometric treatment of chromatographic profile", *Anal. Bioanal. Chem.* **388,** 1859 (2007). doi: 10.1007/s00216-007-1422-9
2. S. Azadmard-Damirchi, "Review of the use of phytosterols as a detection tool for adulteration of olive oil with hazelnut oil", *Food Addit. Contam.* **27,** 1 (2010). doi: 10.1080/02652030903225773
3. S. Mildner-Szkudlarz and H.H. Jeleń, "The potential of different techniques for volatile compounds analysis coupled with PCA for the detection of the adulteration of olive oil with hazelnut oil", *Food Chem.* **110,** 751 (2008). doi: 10.1016/j.foodchem.2008.02.053
4. M. Hajimahmoodi, H.Y. Vander, N. Sadeghi, B. Jannat, M.R. Oveisi and S. Shahbazian, "Gas-chromatographic fatty-acid fingerprints and partial least squares modeling as a basis for the simultaneous determination of edible oil mixtures", *Talanta* **66,** 1108 (2005). doi: 10.1016/j.talanta.2005.01.011
5. G. Flores, M.L. Ruiz Del Castillo, M. Herraiz and G.P. Blanch, "Study of the adulteration of olive oil with hazelnut oil by on-line coupled high performance liquid chromatographic and gas chromatographic analysis of filbertone", *Food Chem.* **97,** 742 (2006). doi: 10.1016/j.foodchem.2005.06.008
6. A.H. El-Hamdy and N.K. El-Fizga, "Detection of olive oil adulteration by measuring its authenticity factor using reversed-phase high performance liquid chromatography", *J. Chromatogr. A.* **708,** 351 (1995). doi: 10.1016/0021-9673(95)00415-J
7. F. Peña, S. Cárdenas, M. Gallego and M. Valcárcel, "Direct olive oil authentication: detection of adulteration of olive oil with hazelnut oil by direct coupling of headspace and mass spectrometry, and multivariate regression techniques", *J. Chromatogr.* **1074,** 215 (2005). doi: 10.1016/j.chroma.2005.03.081
8. V. Baeten, J.A. Fernández Pierna, P. Dardenne, M. Meurens, D.L. García-González and R. Aparicio-Ruiz, "Detection of the presence of hazelnut oil in olive oil by FT-Raman and FT-MIR spectroscopy", *J. Agric. Food. Chem.* **53,** 6201 (2005). doi: 10.1021/jf050595n
9. R. Aparici, M.T. Morales and V. Alonso, "Authentication of European virgin olive oils by their chemical compounds, sensory attributes and consumers attitudes", *J. Agric. Food. Chem.* **45,** 1076 (1997). doi: 10.1021/jf960659h
10. V. Baeten, M. Meurens, M.T. Morales and R. Aparicio, "Detection of virgin olive oil adulteration by Fourier transform Raman spectroscopy", *J. Agric. Food Chem.* **44,** 2225 (1996). doi: 10.1021/jf9600115
11. P. Konstantina, M. George and G. Constantinos, "Synchronous fluorescence spectroscopy for quantitative determination of virgin olive oil adulteration with sunflower oil", *Anal. Bioanal. Chem.* **386,** 1571 (2006). doi: 10.1007/s00216-006-0729-2
12. Y. Rao, B. Xiang, X. Zhou, Z. Wang, S. Xie and J. Xu, "Quantitative and qualitative determination of acid value of peanut oil using near-infrared spectrometry", *J. Food Eng.* **93,** 249 (2009). doi: 10.1016/j.jfoodeng.2009.01.023

13. R.M. El-Abassy, P. Donfack and A.Materny, "Visible Raman spectroscopy for the discrimination of olive oils from different vegetable oils and the detection of adulteration", *J. Raman Spectrosc.* **40,** 1284 (2009). doi: 10.1002/jrs.2279

14. B. Öztürk, A. Ankan and D. Özdemir, "Olive oil adulteration with sunflower and corn oil using molecular fluorescence spectroscopy", in *Olives and Olive Oil in Health and Disease Prevention*, Ed by V.R. Preedy and R.R. Watson. Academic Press, Oxford, UK, p. 451 (2010).

15. G. McLeod, K. Clelland, H. Tapp, E.K. Kemsley, R.H. Wilson, G. Poulter, D. Coombs and C.J. Hewitt, "A comparison of variate pre-selection methods for use in partial least squares regression: a case study on nir spectroscopy applied to monitoring beer fermentation", *J. Food Eng.* **90,** 300 (2009). doi: 10.1016/j.jfoodeng.2008.06.037

16. M. Cruz Sarraguc and J.A. Lopes, "Quality control of pharmaceuticals with NIR: from lab to process line", *Vib. Spectrosc.* **49,** 204 (2009). doi: 10.1016/j.vibspec.2008.07.013

17. L. Xie, X. Ye, D. Liu and Y. Ying, "Quantification of glucose, fructose and sucrose in bayberry juice by NIR and PLS", *Food Chem.* **114,** 1135 (2009). doi: 10.1016/j.foodchem.2008.10.076

18. I.J. Wesley, R.J. Barnes and A.E.J. McGill, "Measurement of adulteration of olive oils by near-infrared spectroscopy", *J. Am. Oil Chem. Soc.* **72,** 289 (1995). doi: 10.1007/BF02541084

19. I.J. Wesley, E. Pacheco and A.E.J. McGill, "Identification of adulterants in olive oils", *J. Am. Oil Chem. Soc.* **73,** 515 (1996). doi: 10.1007/BF02523928

20. H. Yang and J. Irudayaraj, "Comparison of near-infrared, Fourier transform-infrared and Fourier transform-Raman methods for determining olive pomace oil adulteration in extra virgin olive oil", *J. Am. Oil Chem. Soc.* **78,** 889 (2001). doi: 10.1007/s11746-001-0360-6

21. E. Bertran, M. Blanco, J. Coello, H. Iturriaga, S. Maspoch and I. Montoliu, "Near infrared spectrometry and pattern recognition as screening methods for the authentication of virgin olive oils of very close geographical origins", *J. Near Infrared Spectrosc.* **8,** 45 (2000). doi: 10.1255/jnirs.263

22. A.A. Christy, S. Kasemsumran and Y.P. Du, "The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics", *Anal. Sci.* **20,** 935 (2004). doi: 10.2116/analsci.20.935

23. S. Kasemsumran, N. Kang and A. Christy, "Partial least squares processing of near-infrared spectra for discrimination and quantification of adulterated olive oils", *Spectrosc. Lett.* **38,** 839 (2005). doi: 10.1080/00387010500316189

24. J.A. Lopes, P.F. Costa, T.P. Alves, and J.C. Menezes, "Chemometrics in bioprocess engineering: process analytical technology (PAT) applications", *Chemometr. Intell. Lab. Syst.* **74,** 269 (2004). doi: 10.1016/j.chemolab.2004.07.006

25. P. Geladi and B.R. Kowalski, "Partial least squares regression: a tutorial", *Anal. Chim. Acta* **185,** 1 (1986). doi: 10.1016/0003-2670(86)80028-9

26. D.M. Haaland and E.V. Thomas, "Partial least squares methods for spectral analyses. Relation to other quantitative calibration methods and the extraction of qualitative information", *Anal. Chem.* **60,** 1193 (1988). doi: 10.1021/ac00162a020

27. P.D. Wentzell, D.T. Andrews, and B.R. Kowalski, "Maximum likelihood multivariate calibration", *Anal. Chem.* **69,** 2299 (1997). doi: 10.1021/ac961029h

28. C.B. Lucasius, M.L.M. Beckers and G. Kateman, "Genetic algorithms in wavelength selection : a comparative study", *Anal. Chim. Acta* **286,** 135 (1994). doi: 10.1016/0003-2670(94)80155-X

29. U. Hörchner, and J.H. Kalivas, "Further investigation on a comparative study of simulated annealing and genetic algorithm for wavelength selection", *Anal. Chim. Acta* **311,** 1 (1995). doi: 10.1016/0003-2670(95)00163-T

30. R. Leardi, "Application of genetic algorithm–PLS for feature selection in spectral data sets", *J. Chemometr.* **14,** 643 (2000). doi: 10.1002/1099-128X(200009/12)14:5/6<643::AID-CEM621>3.0.CO;2-E

31. H.C. Goicoechea and A.C. Olivieri, "A new family of genetic algorithms for wavelength interval selection in multivariate analytical spectroscopy", *J. Chemometr.* **17,** 338 (2003). doi: 10.1002/cem.812

32. J. Koljonen, T.E.M. Nordling and J.T. Alander, "A review of genetic algorithms in near infrared spectroscopy and chemometrics: past and future", *J. Near Infrared Spectrosc.* **16,** 189 (2008). doi: 10.1255/jnirs.778

33. C. Reynès, S. de Souza, R.T Sabatier, G. Figuères and B. Vidal, "Selection of discriminant wavelength intervals in NIR spectrometry with genetic algorithms", *J. Chemometr.* **20,** 136 (2006). doi: 10.1002/cem.1000

34. D. Özdemir and B. Öztürk, "Genetic multivariate calibration methods for near infrared (NIR) spectroscopic determination of complex mixtures", *Turk. J. Chem.* **28,** 497 (2004).

35. D. Özdemir and B. Öztürk, "Near infrared spectroscopic determination of olive oil adulteration with sunflower and corn oil", *J. Food Drug Anal.* **15,** 40 (2007).

36. B. Uner, İ. Karaman, H. Tanriverdi and D. Özdemir, "Prediction of lignin and extractive content of *Pinus nigra* Arnold. var. Pallasiana tree using near infrared spectroscopy and multivariate calibration", *J. Wood Chem. Technol.* **29,** 24 (2009). doi: 10.1080/02773810802607567